

Two Faces of Desert

Matt King

Philosophical Studies, Vol. 169, No.3: 401-424.

This is the penultimate version. Please cite the published version, which is substantially revised:
<http://link.springer.com/article/10.1007/s11098-013-0188-5>

1. TWO FACES OF RESPONSIBILITY

There are two broadly competing pictures of moral responsibility.¹ On the view I favor, to be responsible for some action is to be related to it in such a way that licenses attributing certain properties to the agent, properties like blameworthiness and praiseworthiness. Responsibility is attributability. The actions we are responsible for reflect distinctive features of our agency, our commitments and values, our judgments about what matters. We are thus “open to appraisal that is therefore appraisal of the individual as an adopter of ends” (229). Frankfurt’s theory of agency is illustrative here. Autonomous and free action, on Frankfurt’s view, reflects an agent’s judgments about the worth of her desires, and is expressive of her evaluative stance toward her motivational scheme.²

A different view understands being responsible in terms of our practices of *holding* each other responsible. Responsibility is accountability, which, “involves a social setting in which we demand (require) certain conduct from one another and respond adversely to one another’s failures to comply with these demands” (229). To hold someone responsible is to hold her to account for her conduct, and involves “...a readiness to respond in certain ways” (235), most notably through the distinctive attitudes constitutive of blame: reactive moral emotions like resentment and indignation. Darwall’s recent view (2006) is representative here. Accountability emphasizes the social context of our responsibility practices,

¹ Here I draw from Watson 1996. Unless otherwise noted, all references to Watson are from that paper. Recently, there have been renewed efforts to scrutinize the general concept of moral responsibility and attempt to draw out various complexities. (See, e.g., Fischer & Tognazzini 2011; Shoemaker 2011; Smith 2012.) But despite this renewed emphasis, there seems no consensus on how to carve up the concept, let alone how to employ the various potential conceptions. Indeed, there is as yet no consensus on which names pick out which positions. So I don’t mean to assume here that Watson’s distinction is either correct or the only possible division. Rather, I use his suggestion as a helpful (and familiar) diagnostic tool (and potentially in ways to which Watson himself would object). The extent to which the two approaches I outline above should be regarded as competitors is part of the aim of the subsequent discussion.

² See Frankfurt 1971.

wherein holding someone responsible requires a certain authority³ to demand specific conduct; indeed, it is to demand that the agent is accountable to us.

My concern in this paper is the relation between moral responsibility and desert. Though not unanimously endorsed,⁴ there is some consensus that the basic concept of moral responsibility is to be understood as responsibility ‘in the desert entailing sense’.⁵ Despite this (cautious) consensus, little has been said about what this notion of desert is that moral responsibility supposedly entails. My aim is to remedy this deficiency for the two faces of responsibility. I argue that each conception calls for a different understanding of desert, and that this places each conception under very different constraints and commitments. By making the connections to desert explicit, we gain a new currency by which to evaluate these extant approaches. To that purpose, I argue that responsibility as accountability carries a significant cost related to desert and praiseworthiness.

2. MORAL RESPONSIBILITY AND DESERT

I want to begin with two exceedingly modest claims. First, moral responsibility is related to blameworthiness and praiseworthiness. We need not take a stand on the precise nature of this relationship. It suffices for my purposes that we glean *something* about moral responsibility if we examine blameworthiness and praiseworthiness. Second, the blameworthy are those that are worthy of blame. This should be an uncontroversial, if unhelpful, claim. This is enough to generate a connection between responsibility and desert in two easy steps. One, agents who are morally responsible for things can be blameworthy or praiseworthy for them. Two, if they are blameworthy or praiseworthy, then they *deserve* blame or praise. So my first proposal is that we understand the worthiness of blame and praise in terms of desert.

This is not a radical proposal. Many theorists characterize moral responsibility as crucially involving desert. For example, Robert Kane claims that “responsibility ascriptions... [have] to do with whether agents deserve blame and praise” (1996: 39). Dana Nelkin’s target is moral responsibility “in the sense that [persons] deserve blame and praise when they act well or badly” (2012: 7). And Galen Strawson writes that “true moral responsibility is responsibility of such a kind that, if we have it, then it *makes sense*, at least, to suppose that it could be just to punish some of us with (eternal) torment in hell and

³ The notion of “authority” is tricky here, since it might connote more than is required to make the point. The idea, however, is that accountability requires those who hold the agent responsible to be in a position which legitimates their demand and reactions.

⁴ See for example McKenna 2009.

⁵ See the next section for claims to the effect that moral responsibility must entail desert, as well as McKenna 2009, pp. 11-12 for a similar assessment and some cautionary comments.

reward others with (eternal) bliss in heaven” (1994: 9).⁶ So there is at least a common thought that moral responsibility involves desert of blame and praise.

It is plausible, then, to suppose that moral responsibility entails a desert claim, and that the blameworthy and praiseworthy deserve blame and praise, respectively. Given moral responsibility’s relation to blameworthiness and praiseworthiness, we can generate the following inference: if S is morally responsible for x, then she deserves blame or praise (in principle).⁷ If this is right, then for any conception of moral responsibility, we can consider its connection to desert. I propose to do just that, examining desert through the lenses of attributability and accountability.

3. DESERT-AS-FITTINGNESS

I cannot provide a *theory* of desert here,⁸ or even a restricted theory of desert entailed by moral responsibility. Instead of this comprehensive task, I seek to outline how being morally responsible for something implies desert for each conception of responsibility. I begin with attributability.

A promising proposal is prompted by Watson’s comments that the relationship between attributability-blame and blameworthiness mirrors the relationship between warranted belief and truth.⁹ I propose to understand the notion of desert at work for attributability by appealing to so-called Fitting-Attitude accounts of value.¹⁰ Fitting-Attitude accounts (FA accounts) analyze a certain range of normative properties in terms of the attitudes it is ‘fitting’ to adopt toward objects that have those properties. So, for instance, the ‘desirable’ is not what we are able to desire, but what is fitting of desire. Similar constructions can be made for linguistically similar terms. Thus, the admirable is what is fitting of admiration, the venerable of veneration, and the contemptible of contempt. FA analyses also extend to other related terms: the scornful is what is fitting of scorn; the fearsome what is fitting of fear. The terms are quite wide-ranging, but all seem to call for an attitude it is fitting to hold toward an object with that property.

⁶ While Strawson doesn’t refer to “desert” explicitly here, it seems implied by talk of “just punishment”. Strawson also explicitly notes that the story about heaven and hell is only meant to illustrate the notion of true moral responsibility at work; certainly the existence of heaven and hell is not actually required.

⁷ The caveat “in principle” is required here and in later formulations because one can be responsible for morally neutral conduct and yet not deserve anything further. Nevertheless, in such cases the right relation supporting desert claims would be present. There are some that deny that morally neutral cases can be instances of *moral* responsibility (see Mele XXXX). Those who favor such a view may omit the caveat.

⁸ Desert is no less troubling a general concept than moral responsibility. See McLeod 2008 for discussion and references.

⁹ As Watson puts it, “[O]ne is worthy of blame just in case the attribution of fault is warranted. ‘S is blameworthy for C’ stands in the same relation to ‘S’s conduct is faulty’ as ‘P’ is true’ stands to ‘P’” (238). This comment is suggestive, though, as is clear below, I don’t fully endorse it.

¹⁰ For a more detailed defense see King 2012.

Most relevant to our purposes, FA analyses appear well-suited to blameworthiness.¹¹ The blameworthy are worthy of blame, those for whom blame is fitting. Considering all of these cases together, we can see that there is a general schema any FA account's analyses will follow:

FA schema: S is phi-able just in case it is 'fitting' to phi S.

It is important to note that 'fitting' here is a placeholder for whatever the right normative relation is. So, the phi-able may be what ought to be phi-d, or is correct to phi, or is rational to phi, or is fill-in-the-blank to phi. Different FA accounts may use a different relation, provided it is a normative relation.¹² My proposal is that we use 'desert' in this place: thus, desert-as-fittingness. Placing desert in the schema, we get the following proposal for the attributability conception of blameworthiness:

Att-BW: S is blameworthy just in case S deserves blame.

By itself, this proposal may look uninteresting. But I want to highlight several significant features. First, it secures, at least on the surface, the desert entailing aspect of moral responsibility. To be morally responsible entails a desert claim, and desert-as-fittingness gets us some way toward a justification for why. If desert is a good way to understand the notion of fittingness in FA analyses,¹³ and if an FA analysis is a promising general approach to analyzing terms like blameworthiness,¹⁴ then it will follow that being blameworthy yields a claim about desert.

Moreover, if we put this proposal together with the idea that attributability can serve as a *ground* for desert claims, we see how each can help illuminate the other. The blameworthy are those that deserve blame, and they deserve it in virtue of something which is attributable to them. The basis for the desert claim is the attributability relation, which makes clear how desert is entailed (i.e., in virtue of satisfying attributability conditions). I don't claim that this is the only possible positive view about desert to give for attributability; much less that desert-as-fittingness *follows* from an understanding of attributability. But I think these attractions show it to merit consideration.

Additionally, desert-as-fittingness helps vindicate Watson's defense of attributability as a "deep" assessment of individuals (231-234). Wolf (1990) argues that attributability is at best a superficial assessment, applying equally well to the clumsy or unskilled as to the blameworthy. Watson's reply is that as attributability-blame targets elements of one's agency, the agent's values and commitments expressed

¹¹ It is worth noting that the similarity in a suffix is not sufficient to suggest an additional term well-suited to FA analysis; one must also have the suggestion of an attitude. For example, 'seaworthy' refers to a craft suitable for sea travel. Similarly, the edible *is* what we are able to eat. Neither term contains reference to an attitude, despite the similarity in suffix to FA-ready terms. Compare, however, 'noteworthy' or 'trustworthy'.

¹² This isn't to say any characterization of fittingness is as good as another. But we need not evaluate between respective proposals here.

¹³ See King 2012.

¹⁴ We do have reason to think that FA analyses are very promising. See Schroeder 2010 for a defense of this optimism.

through action, it targets not superficial features of persons but rather the core of our practical identities (234). This is in part what makes it appropriate to talk of attributability as a form of *moral* responsibility.

One might worry that since FA analyses do not apply exclusively to moral terms, desert-as-fittingness only obscures Watson's observation. After all, FA analyses apply to properties like the fearsome and disgusting. But to judge something fearsome or disgusting is not (ordinarily) to make a deep assessment, and it certainly doesn't reference its practical identity. Nevertheless, desert-as-fittingness has resources to accommodate this thought.¹⁵ Consider the differences among the attitudes of blame, fear, and disgust. Blame typically targets features of agency at least when used in evaluative, rather than merely explanatory, contexts.¹⁶ Fear and disgust, on the other hand, often target purely superficial qualities, such as an object's dangerous or nauseating features.¹⁷ So, one way to capture the requisite depth is by reference to the attitude deserved. This corroborates Watson's claim that attributability concerns aspects of an agent's practical identity. It also vindicates attributability-blame as a moral attitude. As Strawson (1962) notes, we care about the attitudes of others in a distinctive way; their values and commitments carry an importance to our interpersonal relations and our lives as persons that non-moral attitudes lack. Desert-as-fittingness locates the distinctively moral sense of attributability within the relevance of blame to moral agents, and the sorts of things attributability-blame targets (e.g., values, commitments).¹⁸

Still, one might wonder whether desert-as-fittingness is an appropriate conception of desert. How can the fearsome be said to deserve fear as the blameworthy deserve blame? But it is important to keep in mind that while the desert relation being appealed to here is the same, the desert *basis* is not. According to our initial assumption, all that is required is that moral responsibility entails desert; the blameworthy deserve blame in virtue of being morally responsible. And desert-as-fittingness maintains and vindicates those claims. Presumably, the fearsome will not deserve fear because they are morally responsible for anything, at least not in ordinary contexts. To say that they nevertheless deserve fear does no injustice to our concept of desert. It is widely recognized that we make desert claims about non-persons, who arguably cannot be responsible for anything. That the Statue of Liberty deserves preservation or that a species deserves protection does not imply anything about moral responsibility. This is because the entailment is one-way only. Moral responsibility entails something about desert; as I have suggested, it states the basis for a desert claim. But such claims do not go through the other way around.

¹⁵ The defense that follows draws from King 2012.

¹⁶ The faulty strut may be 'to blame' for the bridge's collapse, but this isn't of any moral relevance. Indeed, to say that it was faulty and to blame isn't really an evaluation. We could equally well say that the cause of the collapse was a weakened strut.

¹⁷ Though we can, of course, be 'disgusted' by someone's wrongful behavior or viciousness. Still, even were we to distinguish between moral disgust and non-moral disgust completely, an FA analysis of the latter might nevertheless be correct.

¹⁸ Of course, another strand of Strawson's thought might be that the reactive attitudes (like resentment, and indignation) presuppose *holding* their target accountable. Strawson has certainly been interpreted this way (see Wallace 1994). But whether or not this is the best exegesis of Strawson's view is orthogonal to the discussion here, for we could just as easily develop a Strawsonian view along similar, but different, lines, wherein the reactive attitudes only presuppose their targets *to be* responsible. If the presupposition is only that the impetus for the reactive attitude is attributable to the agent, and therefore reflects the agent's values and commitments, then the resulting view can still make use of Strawson's basic observation, even if Strawson himself (or his interpreters) took the view in a different direction.

Desert-as-fittingness helps make clear an important feature of the relationship between attributability-blame and blameworthiness. As blame is deserved only by the blameworthy, and as attributability-blame targets the features of practical identity, the sorts of reasons relevant to who deserves what on desert-as-fittingness would seemingly be limited to considerations regarding the agent herself. That something was an accident undermines the claim that it reveals anything about an agent's practical identity; in a similar fashion, so can the claim that it was a mistake. But both these features point to the agent's intentions and beliefs. If the agent was drugged or suffers from spasms, these too serve to sever the connection between actions and values. Desert-as-fittingness demands that blame fit its object, and it fits its object so long as its object is blameworthy for *x*, which requires that *x* be attributable to it.¹⁹

Desert-as-fittingness, then, not only preserves a general thought about desert, that the desert basis must be a fact 'about' that thing;²⁰ it also seems to suit a particular conception of desert in the context of moral responsibility. Derk Pereboom claims that a theory of moral responsibility ought to give the conditions satisfaction of which would entail that the agent "would deserve blame if he understood that it was morally wrong" (2007: 197). For him, basic desert "is basic in the sense that the agent... would deserve the blame... just by virtue of having performed the action, and not, for example, by way of consequentialist considerations" (2001: xx) nor by way of a contractualist account (2009: 22). I think this characterization is still imprecise. Surely one doesn't deserve blame simply for having performed an action. Rather, it must be for performing an action in a particular way, or through engaging certain capacities or powers.²¹ Nevertheless, desert-as-fittingness looks consistent with Pereboom's characterization of desert; and attributability helps formulate what the relevant capacities or "way" of acting might be (i.e., self-disclosing ones).²²

4. FITTINGNESS AND ACCOUNTABILITY

I want to pause at this point to argue that desert-as-fittingness cannot be usefully extended to accountability. To show this, we will need to first recall how accountability-blame differs from the attributability variety.

Accountability takes its cue from our practices of *holding* each other responsible, and thus accountability-blame involves a distinctive stance. This stance, Watson tells us, is to require or demand certain conduct from others, which is to "lay it down that that unless the agent so behaves she will be liable to certain adverse or unwelcome treatment" (236-7). Thus, Watson connects 'accountability' essentially with 'liability to sanctions'. Accountability-blame is therefore following through with those sanctions. But his discussion of sanctions and adverse treatment is ambiguous between two possibilities.

¹⁹ With the same caveat as before, that *x* be negatively evaluable in some way.

²⁰ Feinberg 1970. See also, McLeod 2008.

²¹ I think Pereboom would agree on this score. He cannot doubt that determined agents still perform actions, though they don't count as deserving anything.

²² Of course, Pereboom would not concede that any set of compatibilist conditions would be sufficient for grounding basic desert. But that is a separate issue.

On the one hand, it may be sufficient to target the blameworthy with the so-called reactive attitudes of resentment and indignation.²³ Following Glover (1970), Watson notes that it is disagreeable to have others resent you, and he holds that the sanctions involved in our accountability practices need be no more than disagreeable attitudes. But this raises two concerns. One, resentment that was not expressed might very well not be disagreeable, nor involve the meaningful social context that accountability presumably draws upon. Private blame can only be disagreeable if the target is aware of it. And while I can be taken to hold you responsible even if I only resent you in private, there is still a sense in which I am not actively holding you responsible to myself or others, which might reasonably involve at least something of an exchange. This is why those elaborating upon accountability often appeal to the notion of a moral conversation.²⁴

The second concern is that, to the extent that we allow for unexpressed blame to count as sanctions, we blur the distinction between attributability-blame and accountability-blame.²⁵ As I have argued, attributability-blame plausibly is not limited to mere judgments of blameworthiness, but also can involve blaming the subject by means of taking up reactive attitudes like resentment and indignation. To take up these attitudes is consistent with not holding those agents accountable, for it needn't invoke social contexts, distinctively interpersonal demands, or associated concepts like authority or liability. If accountability isn't distinguished by involving these related notions essentially, I think we begin to lose our grip on what were supposed to be two distinctive perspectives on moral responsibility. It is for these concerns that I think it best to construe accountability blame as the *imposition* of sanctions, which requires the expression of reactive attitudes or associated treatments, like punishment.²⁶

If accountability-blame is best understood as involving the imposition of sanctions, then this goes beyond the simple holding of attitudes. As such, it is unclear immediately how best to apply the model of fitting-*attitude* accounts of value, as we did with attributability-blame. Our FA schema for blameworthiness held that the blameworthy are those it is fitting to blame. I suggested that we understand that blame in terms of the reactive attitudes, like resentment and indignation, and that fittingness is to be understood in terms of desert, such that the blameworthy are those that deserve resentment or indignation. If desert-as-fittingness can be applied here, though, it will have to proceed a bit differently. For FA accounts, while accounting for the fittingness of attitudes, remain silent on what may be *further* warranted. So if accountability-blame is to be understood as the imposition of sanctions, FA accounts aren't a suitable model without modification.

²³ See Strawson 1962 and Wallace 1994.

²⁴ This is a major component of Watson's (1987) detailed and insightful discussion of the case of Robert Alton Harris. The significance of (or the metaphor of) conversation is also central to McKenna 2012.

²⁵ Watson goes so far as to say that, should private judgments of blame constitute harms, this would make "sanctions inseparable from aretaic judgments" (247, n.41). Watson seems to favor withholding meaningful blame from attributability altogether, whereas I favor making accountability concern expressions of blame alone. My reasons are in the text. Fischer and Tognazzini (2011: 3 and n.7) make a similar restriction as Watson's.

²⁶ All I mean here is that punishment can be a blame-related treatment. I don't think it is necessarily so. It is perfectly consistent to think the blameworthy deserve blame but to deny that they deserve punishment, on the grounds that no one deserves punishment. One's theory of punishment is not conceptually beholden to one's theory of moral responsibility.

Nonetheless, perhaps modifications are available. While FA accounts have been popular in the theory of value, we needn't limit ourselves to the fittingness of attitudes alone. We might expand our purview to include states of affairs. In doing so, we would be forced to give up part of what has been attractive about FA accounts in general; namely, that suitability of an FA account is prompted by a linguistic connection between, e.g., blameworthiness and the attitude marked by 'blame' (and likewise for other plausible candidates, like the admirable or fearsome). But desert-as-fittingness needn't be so narrowly construed, at least if we're interested in trying to make good its extension to accountability.

If this is our aim, then we can draw upon some simple observations where desert claims appear to depend upon a notion of 'fit'. Consider retributive theories of punishment, wherein desert plays a defining role. The punishment deserved is the one that "fits" the crime.²⁷ Moreover, we know that this notion of fit is meant to capture something different than literally repaying the criminal in kind. There's a sense of restoring a balance, of making things right. There is something about retribution that invites a sense of making punishment fitting, whether or not fittingness is part of the definition of retributive punishment. The phrase 'getting one's just deserts' also suggests such fittingness, as when a villain is hoist with his own petard. I think it plausible to suppose that a large set of desert claims involves achieving a certain state of affairs which is most fitting to the circumstances. Punishment is just one possible instantiation of such a state of affairs.²⁸

Similarly, when we talk about competitions, we can readily distinguish winners from those who deserved to win. And when those who most deserve to win do in fact win, there often seems to be a special fittingness to that state of affairs. The underdog who worked hardest and overcame the most is a most fitting champion. We also talk of people deserving each other. This can be either a tribute to the relationship or a derisive slight. But in both cases we are appealing at least to a certain sort of fit.

If these examples are on target, then we have some motivation for extending use of fittingness beyond attitudes and their objects to states of affairs as well. While this modification is consistent with desert-as-fittingness, insofar as it retains an intuitive relation between desert and fittingness, it still undermines somewhat the extension from attributability. We seem unable to simply translate desert-as-fittingness into accountability-terms. Part of the difficulty, is that FA accounts carry with them the idea of a unified normativity. Whatever the 'fitting' relation for all FA analyzable terms, it is the same and sensitive in the same way to reasons. Thus, desert-as-fittingness could draw on that generality and assert its bold conjecture: that desert is to be understood *as* the fittingness relation in FA accounts. Expanding desert-as-fittingness to concern states of affairs instead loses much of what made the model distinctive in the first place.

A second part of the difficulty concerns not what is deserved (attitudes vs. expressed treatment), but what is relevant to whether it is deserved. Desert-as-fittingness for attributability seemed to partially vindicate the contention that desert is grounded on facts about the subject itself. But accountability-blame makes relevant considerations about those holding the subject accountable. Thus, if I lack the standing to hold you to a legitimate demand, then you cannot be accountable to me for violating that demand. This

²⁷ As enshrined in the notion of *lex talionis*. All I'm noting here is the connection between desert and fittingness retributivism often invokes; I am not defending retributivism.

²⁸ And, indeed, the 'fittingness' may dictate that only some among a set of potential punishers are fitting actual punishers.

makes the standing of others, their ‘authority, relevant to the fittingness of accountability-blame. It may be more fitting for certain individuals to impose the sanctions (e.g., those wronged); likewise, it may be especially *unfitting* for certain individuals to impose the sanctions (e.g., those who did the exact same thing).²⁹

But the standing of the appraisers doesn’t bear on whether attitudes fit their objects. Take a non-moral example. Whether something is fearsome does not depend on my standing as an appraiser; it depends on something like how dangerous that thing is.³⁰ Of course, certain things may be dangerous to me and not to others. A bear is dangerous to me, but not to Superman. Likewise, kryptonite is dangerous to Superman, but not to me. But these facts still concern the thing in question, its ability to hurt or damage others. Most importantly, they don’t depend upon my status as a fearer, or upon my past as a fearer, or my own fearsomeness. Fear is no less fitting of poison for one’s being impressively dangerous oneself. So, though Superman might be fearsome (to a petty criminal), this does not affect kryptonite’s fearsomeness (for him). If I am especially frightful, fearing many things, this does not affect my standing vis-à-vis the fearsome; it is still fitting to fear whatever is fearsome, no matter my past fearings (appropriate or not) or tendency to fear. But this isn’t obviously true for accountability-blame. If I am morally vicious, I may therein lose my standing to legitimately hold others to account for their wrongdoing.³¹ If I am quick to blame, blaming all manner of people for all manner of faults and non-faults, then this may also diminish my standing.

So, while both models might warrant the name ‘desert-as-fittingness’, if nothing of actual substance has been carried over between attributability and accountability, then there may be no one desert-as-fittingness view to suffice as a model for *both* perspectives. And as reflection on the standing of appraisers reveals, we may indeed require more than ‘fittingness-desert’ to successfully develop an accountability view.

5. DESERT-AS-FAIRNESS

In contrast to attributability, being accountable for x “is a three-term relationship in which one individual or group is held by another to certain expectations or demands or requirements,” and “involves a

²⁹ For a more detailed discussion of kinds of cases that might be relevant here, see Smith 2007.

³⁰ Here I am following, for illustrative purposes, a view similar to D’Arms and Jacobsen 2001. There is reasonable disagreement about how to fill out the details of FA accounts. My preference is that the attitudes target certain features which the object of those attitudes must then possess. So, for attributability, those features might be spelled out by the conditions adopted by self-disclosure views. Blame targets aspects of an agent’s will. Fear targets those features that are, say, dangerous. Not everyone agrees this is how FA accounts should proceed. Despite the differences in how we work out the details, however, there is rough agreement that certain sorts of reasons are to be ruled out as “reasons of the wrong kind” (e.g., that you would suffer horrible pain should you fail to fear, or blame, or admire). My suggestion here is the considerations that are essential to understanding accountability will plausibly fall into this category. See Schroeder 2010 for a partial comparison of approaches to the wrong kind of reasons problem.

³¹ And if not lost, my standing appears at least diminished.

readiness to respond in certain ways” should those demands go unmet (235).³² As noted above, the making of demands or requirements and the readiness to respond, especially in negative ways, makes accountability-blame look like liability to sanctions. Characterized in this way, accountability-blame suggests a connection to both authority and considerations of fairness. We might reasonably ask, as Watson does, where the authority to issue these demands comes from and under what conditions it is fair to respond with sanctions.

Watson doesn’t say much about the authority to make moral demands, beyond noting clear cases in which it wouldn’t obtain.³³ Darwall (2006), however, connects this authority to the moral value of respect for persons. As an admittedly skeletal summary, respect for persons grounds an interpersonal authority to demand compliance with obligations. This can be seen clearly in promissory obligations, wherein the making of a promise authorizes the promisee to hold the promiser to the promise’s content.³⁴ Though promising is a special case, we can see accountability being naturally paired with contractualist approaches in ethics, where holding each other to shared obligations and responsibilities with the authority to demand compliance is front and center.³⁵

Failure to comply with legitimate demands opens the accountable to legitimate sanctions. The fact that accountability-blame is sanction-like makes considerations of fairness particularly relevant. Indeed, it introduces concerns about avoidability, provided it seems plausible that “[i]t is unfair to impose sanctions upon people unless they had a reasonable opportunity to avoid incurring them” (237). That accountability is essentially an interpersonal notion means that the deservingness of desert is to be understood within that social context, understood in terms of our interpersonal demands, the authority to make them, and the authority to respond in the sanction-like ways distinctive of holding others responsible.

If all this right, then we can say something positive about the desert relation for accountability-blame. We might understand the desert involved in terms of fairness. The blameworthy deserve blame (for x) because they are accountable (for x) such that it is fair to blame them (for x). Talk about desert is often talk about fairness; it can be unfair to be denied what we deserve just as it can be unfair to be given what we do not deserve. Desert-as-fairness draws on these connections between desert and fairness. Given accountability-blame’s social context, the desert of blame is taken to be (at least partially) a social conception. To hold others accountable is not just to adopt a stance where one expects certain demands to be met and responds negatively when they are flouted. Adopting that stance is beholden to considerations of fairness in two ways. First, it must be fair to impose the demand in the first place. These will be considerations concerning what our legitimate obligations are, at least in terms of what others may

³² Watson notes and sets aside the important self-reflexive case where one holds oneself accountable. He also discusses interesting non-moral instances, like a janitor being held accountable for neglecting the furnace (236). Like Watson, I will focus on the moral, non-self-reflexive cases.

³³ His example is a hijacker making one of the hostages “accountable” for the behavior of the others, on pain of death. Plausibly, the hijacker lacks the standing to legitimately make such a demand (237).

³⁴ Watson also notes the paradigmatic nature of promises for highlighting authority in the context of accountability (237-8).

³⁵ Indeed, Darwall’s (2006) approach is best understood as proceeding from a contractualist circle of interdefined moral notions: obligation, authority, respect, and accountability.

legitimately demand of us. So, for instance, the dictum that ‘ought implies can’ may be particularly relevant here.³⁶ If it is unfair to demand some conduct from agents, it is unfair to blame them for failing the demand.

Second, the resulting sanction-like ways of responding must be fair. There are two different kinds of considerations that could make blaming treatment unfair. One kind of consideration concerns the blamed agent as he is. So, it can be unfair to blame me for something I haven’t done, or something I had no control over and no reasonable opportunity to affect or avoid, such as a hiccup or spasm. What is at issue here is whether it would be fair to blame anyone for such things.³⁷ These are what we might call considerations of universal fairness. In contrast, another kind of considerations concerns the context between persons, often between the blamers and the one blamed. So, it might be unfair for us to blame Rick, if we aren’t also blaming two Steves, who did the very same thing. Or it might be unfair for me to blame Rick, given that I did the very same thing this morning. These are what we might call considerations of comparative fairness.

While invoking fairness may help to draw accountability closer to related moral concepts like obligation and authority, it also expands the circle of considerations relevant to the responsibility relation. Accountability is given shape by the notion of a moral community, those who are making the demands and to whom one is accountable. Their inclusion in the accountability relation makes deserved accountability-blame sensitive to more than an attitude fitting its object. It suggests blame as being constituted by a form of treatment with a certain force, which makes the very expression of blame a moral matter, sensitive to moral considerations.³⁸ This is in contrast to the more restricted picture presented by attributability.

6. FAIRNESS AND ATTRIBUTABILITY

The preceding already hints at the difficulties for extending desert-as-fairness from accountability to attributability. For accountability-blame, the focus is on the relation between the blameworthy subject and the moral community enforcing the shared demands. It is plausible to think we have a good sense of how our practices of holding someone accountable or imposing sanctions are sensitive to fairness.³⁹ Watson’s discussion treats accountability as a “legal-like practice, an informal institution serving the ends of social

³⁶ Watson takes something of a stand on this point, arguing that one can be legitimately held accountable for the behavior of others (against some standard) even if one could not have ensured their behavior met the standard. His example is hiring someone to keep order in a dance hall, where blaming them for the disorder is fair even though they may not have been able to control the crowd, so long as they agreed to the terms and conditions. This may be a special case, as it involves an explicit agreement, and so still leaves open general questions about the demandingness of obligation in general (237).

³⁷ It is perhaps worth noting that these considerations also tend to disrupt the connection between one’s action and their values and commitments. That is, they also plausibly affect attributability.

³⁸ For a critical discussion of the relevance of fairness to blame, brought out through a discussion of its force, see Hieronymi 2004.

³⁹ Compare, Gideon Rosen: While “[i]t is obviously not a matter of equal distribution of the reactive attitudes...or of procedural impartiality in their application”, nevertheless, “we possess a robust body of opinion about when it is fair to treat someone adversely for what he has done” (2002: 74).

regulation and/or retributive and compensatory justice” (239). This makes reasons which concern the blamers, those in the moral community to whom the blameworthy are accountable, obviously relevant to the fairness of accountability-blame, but not obviously relevant to deserved attributability-blame.⁴⁰

This can perhaps be most clearly brought out by focusing on considerations of comparative fairness. Suppose Hal has stolen Jordan’s car.⁴¹ This plausibly violates an obligation it is reasonable to demand Hal comply with, and so accountability-blame is in principle warranted. But suppose that Jordan is himself an unrepentant car thief. Here his blame looks hypocritical; who is Jordan to blame Hal for something that Jordan does regularly? If such blame is unfair, then by accountability’s standards Hal may not deserve accountability-blame from Jordan for stealing the car.

Such examples help to make the needed point regarding comparative fairness. Of course, one might think such examples are too idiosyncratic. Maybe Jordan’s blame is unfair because hypocritical, but this doesn’t make it unfair for Carol, Jordan’s wife, to blame Hal, as Carol is not a car thief. Still, there are plausible examples of general features which might make most of those in a moral community unfair blamers. Angela Smith (2007) notes that it may be fair only for certain individuals in special relationships to blame each other; those outside of these relationship are in no position to blame. We can imagine here the special place that siblings or close friends hold which may license them a special authority not only to demand certain conduct but also to impose the requisite sanctions. Notably, this can be true even for general violations. Someone who manifests demeaning attitudes towards his wife at a party may be unrepachable by *me*; while a good friend may pull him aside and reprimand fairly, I cannot do so.⁴² There are still further cases that may show that no one in the moral community can fairly blame. Given concerns about circumstantial luck, it may be that none of us are in position to blame Nazis or Attila the Hun’s armies, for if we had been in similar circumstances, growing up German in the 1920s or born a Hun, we too would have behaved the same way.

I don’t mean to suggest that in all of these cases the blame would necessarily be unfair, nor that a view about accountability is committed to holding that it would be. For my purposes here it is enough for the cases to show that the fairness of accountability-blame, and thereby what is deserved according to desert-as-fairness, is sensitive to certain reasons that don’t appear relevant for the desert of attributability-blame. There can be little doubt that the car theft is attributable to Hal, or that the wife’s poor treatment is attributable to the husband, or that Nazi and Hun atrocities alike are attributable to their respective perpetrators. In light of this, attributability-blame appears deserved, and not obviously unfair.

It remains unclear, however, what could make attributability-blame unfair in any *distinctive* sense. Of course, if one is blamed for something that is not even attributable to them, this *might* count as unfair in some restricted sense. But it isn’t obvious that we can get even this much. Recall that considerations of fairness were naturally raised by our practices of accountability because accountability-blame essentially involved expressions of blame or exposure to sanctioning treatment. Subjection to sanctions makes fairness crucially relevant. This element is largely absent from attributions of blame. Consider an

⁴⁰ Again, see Smith 2007 for related discussion, though her focus is not on desert.

⁴¹ I owe the example to Pete Graham.

⁴² Smith 2007: 478-9.

analogous case, in which Venus thinks beautiful that which is really ugly.⁴³ While unwarranted (and perhaps *unfitting*), it isn't clear that anything unfair has been done. And this seems right even if we take the target of the thought to be a person (rather than, say, a piece of art). Similarly, even being wrong about the morally faulty isn't obviously to be unfair. If Venus thinks the decent is really vicious, if she is privately indignant, this isn't to be fair or unfair. Concerns about fairness just don't get a grip here.

Given that we have been concerned with distinguishing between attributability and accountability, it should not be surprising that one can have something attributable to them that one is not accountable to others for. Watson makes this plain. But what has been suggested by the foregoing discussion is that, say, the propriety of remonstrating the husband at the party is relevant to what he deserves accountability-wise, but not what he deserves attributability-wise. Thus, the differences between the two conceptions of responsibility suggest a continued difference for desert.

7. PRAISEWORTHINESS AND DESERT

For both conceptions of desert, we have seen each fit one of the two faces of responsibility fairly well, but that each must contend with certain difficulties when being extended to the other. Desert-as-fittingness appears apt for attributability, but not accountability. Desert-as-fairness is well-suited to accountability, but not attributability. So far, then, reflections on desert may create more space between the two faces than initially thought. In this section, I consider another feature that threatens to widen the gap further.

For the most part, I've stuck with talking about blame and blameworthiness. But our original point of departure, that moral responsibility entails desert, was meant to be more general than the discussion has perhaps suggested. When Pereboom insists that he's interested in "basic desert", he means not just that the conditions on moral responsibility will justify assessments of blameworthiness, but of praiseworthiness, too. Indeed, pretheoretically, blameworthiness and praiseworthiness seem intimately related; of those who are morally responsible, the blameworthy are worse than the praiseworthy. And just as the blameworthy deserve blame, the praiseworthy deserve praise. So desert isn't just relevant to moral responsibility and blameworthiness. It is also relevant to praiseworthiness.

We would do well, then, to consider both perspectives on responsibility from the angle of deserved praise. Immediately, this poses something of a problem. It isn't clear how to make sense of accountability-praise. Attributability-praise is easy enough. Objects attributable to the agent express or reveal that agent's values and commitments. But nothing in this characterization specifies whether those values and commitments are to be negatively or positively evaluated. There appears to be no obstacle to running the same account through for attributability-praise that we did for attributability-blame. Attributability-praise concerns the moral excellences and virtues of character expressed in action, and the praiseworthy deserve praise in virtue of satisfying the same self-disclosing conditions as concern attributability generally.

But accountability is harder to sort out. Consider its typical characterizations. When I first introduced it, I quoted Watson: accountability "involves a social setting in which we demand (require) certain conduct from one another and respond adversely to one another's failures to comply with these

⁴³ Watson also compares aretaic moral judgments with aesthetic ones.

demands” (229). Accountability draws from our practices of holding each other responsible, which, unlike holding others to *be* responsible, commits us to a stance of expectation and demand, one that invites the imposition of sanctions (236).⁴⁴ None of this looks remotely like praise. Of course, this may just betray our preoccupation with blame and negative evaluations. As Watson suggests, this is perhaps to be expected given that blame can be a more serious affair (242).

Watson agrees that there is no straightforward positive counterpart to accountability-blame. Accountability as liability to sanctions doesn’t translate well to talk of praise. Nevertheless, Watson thinks we do have a counterpart notion in the way of bestowing awards for exemplary conduct (242). Just as blame is a response to the violation of demands, praise is recognition of respecting or exceeding our demands. I take no issue with the claim that praise is a response to good deeds. But I don’t think Watson has identified a counterpart notion to accountability-blame. And this isn’t Watson’s failing; rather, this should be the expected result, for I don’t think accountability can be sensibly extended to cover praise.

When we hold someone accountable to some demand, this involves more than just the claim that we hold them to demands in general. To be a moral agent means that you are subject to obligations. Thus, to treat someone as a moral agent is to hold them to their obligations. But to hold someone accountable, to accountability-blame them, is to hold them to a particular demand, the one they violated. In the accountability-praise case, given that we are dealing with a moral agent, it still makes sense to say we are holding them to moral demands. But in praising the agent, I am not *thereby* holding her to any demand. These look to be separate claims. Note that by saying that you hold so-and-so responsible for x looks tantamount to blaming them; it would be an odd expression indeed to imply praise. If Sherry has donated money to Oxfam, Gina doesn’t hold her accountable for it. There’s nothing to ‘account’ for. Gina may, of course, hold Sherry accountable in general, for Sherry is a moral agent. But in praising her act of generosity, she is not holding her to a demand to be generous. Moreover, suppose there are no positive duties to aid. This wouldn’t show that helping others is not praiseworthy. So, Sherry can be praiseworthy for donating to Oxfam even if there is no obligation to do so, and therefore, plausibly, no legitimate demand we can make of her to do so.⁴⁵

I think we should want a sensible account of praiseworthiness from our conceptions of moral responsibility, one that parallels our account of blameworthiness. I remain unconvinced that accountability can provide such an account.⁴⁶ As I developed it, the conception for accountability was desert-as-fairness. Suppose we try understanding praise in terms of the fairness of rewards, in the same way that blame can be understood in terms of the fairness of sanctions.⁴⁷ If I blame Fred for a wholly

⁴⁴ Smith 2007 also emphasizes the difference between holding others responsible and holding them to be responsible. The latter use of ‘hold’ is in the sense of entertaining a proposition, rather than an interpersonal stance involving demands and distinctive responses.

⁴⁵ Notice that the same cannot be said of blameworthiness. If there were no negative obligations, we plausibly couldn’t be blameworthy for anything. This is the thought behind the so-called Deontic Argument. No matter what one thinks of that argument in general, if there were no negative obligations, accountability-blame would be impossible, for there would be no legitimate demands to hold anyone to.

⁴⁶ This may be why, in the two most developed presentations of accountability theories, Darwall 2006 never so much as mentions praiseworthiness, and Wallace 1994 explicitly sets it to one side.

⁴⁷ This is something Watson suggests (242).

accidental shooting,⁴⁸ then it seems I've treated him unfairly. But such an account clearly will not work for praise. If I praise Wayne for, say, saving a child's life by accident,⁴⁹ it isn't clear that I've done anything unfair to *him*. In benefiting the agent with undeserved praise, I don't do anything unfair to the agent. A distribution of rewards itself might be unfair, if we assume that my giving praise to the unworthy agent deprives some other agent who is more deserving. But such a distribution would be unfair precisely to those whom I'm *not* praising. In this way, it clearly differs from accountability-blame with respect to what is *deserved*.

This is not to say that fairness has no relevance to discussions of praise. While praising the undeserving may not be unfair, certain comparative treatments might be. I might be unfair in praising Wayne if I *fail* to praise Bruce, who has saved a child's life intentionally. But what's unfair here isn't praising Wayne, but rather not praising Bruce. And if I happen to praise them both, it isn't clear that I've done anything unfair to anyone. One who gets what he deserves has little complaint against giving someone else what they don't deserve. And, of course, in the case as given, where there is no Bruce, praising Wayne for his accidental saving cannot plausibly be denying anyone else deserved praise. So praise doesn't seem to be limited in the way necessary to raise concerns about distributive justice. Indeed, this may make construals of praise along the lines of 'reward' very misleading.

Here is what I think accounts for the difference. Accountability-blame and -praise involve different comparative commitments. For blame, we start with no one being blamed for anything, and any deviation must be justified as fair. Thus, only deserved blame can be fair; blaming the undeserving is to subject them to undeserved adverse treatment. But the comparative baseline for praise is different. Praising the undeserving is not to *subject* them to anything, thus there is nothing that demands independent justification. What calls for justifying is the failure to praise the praiseworthy (where such failures might be made particularly salient in light of some instance of undeserved praise). This is, on its face, a different matter than justifying the imposition of sanctions, which is plausibly required for any instance of accountability-blame (even if there were only one blameworthy agent in existence).⁵⁰ Instead, failing to praise the praiseworthy is a failure to give them what they deserve, but this doesn't make additional praise heaped on the undeserving unfair to anyone. And even if it were, it would be unfair to precisely those we are not praising, and we would be asked to justify what we are not doing (praising the praiseworthy), in sharp contrast to how fairness is necessary to justified accountability-blame.

Forgoing accountability-praise may satisfy those wedded to views of responsibility as accountability, but I think it is a significant cost to bear. Desert-as-fittingness implies that the blameworthy and praiseworthy deserve blame and praise, respectively, and for the same reasons. Self-disclosure

⁴⁸ We might suppose that Fred shoots Barney in the eye with a cork that has "shot" off from the champagne bottle Fred was opening, ricocheting off walls and other objects, only to hit Barney as he unexpectedly enters the kitchen.

⁴⁹ This is slightly ambiguous. One might mean that Wayne saved the child though he was trying not to. I obviously mean the other interpretation: that he was trying to do something else entirely, but wound up saving a child's life (though he didn't know at the time he was doing so).

⁵⁰ This possibility suggests an intriguing line of inquiry, however. Though guilt is thought by some to be first-personal blame, it is less clear that it can count as unfair. Can individuals *subject* themselves to sanctions? And if they do so undeservedly, can they be treating *themselves* unfairly? I think these questions are provocative, and their answers are important for both the accountability perspective of responsibility and our understanding of guilt. Still, I cannot pursue them here.

conditions are ecumenical with respect to virtuous and vicious conduct, and there is no reason to think that the desert-as-fittingness schema can't be applied to both. In contrast, undeserved blame is clearly unfair treatment of the agent, whereas undeserved praise is not. Thus, deserved blame must seemingly be treated differently from deserved praise given desert-as-fairness. Since neither conception of desert can be extended to the other face of responsibility, we have two diverging positions not only with respect to moral responsibility, but with respect to desert as well. And since responsibility as accountability carries a significant cost with respect to desert, this gives us some reason to favor responsibility as attributability.

A fully adequate argument in favor of attributability would show that its benefits not only outweigh its costs, but outweigh a similar calculation for accountability. So my conclusion is more restricted than an outright endorsement of attributability as the justified conception of moral responsibility. Instead, I only conclude that reflections on desert widens the gap between the two faces of responsibility, and that attributability carries the more plausible commitments with respect to deserving blame and praise. I consider this a significant virtue of that approach.⁵¹

REFERENCES

- Brandt, R. (1958). "Blameworthiness and Obligation." In Melden, A. (ed.), *Essays in Moral Philosophy*. University of Washington Press.
- D'Arms, J. and D. Jacobsen. (2000). "The Moralistic Fallacy: On the "Appropriateness" of Emotions." *Philosophy and Phenomenological Research* 61: 65-90.
- Darwall, S. (2006). *The Second-Person Standpoint: Respect, Morality, and Accountability*. Harvard University Press.
- Feinberg, J. (1970). "Justice and Personal Desert." In Feinberg, J. *Doing and Deserving*. Princeton University Press.
- Fischer, J.M. and M. Ravizza. (1998). *Responsibility and Control*. Cambridge University Press.
- Fisher, J.M. and N. Tognazzini. (2011). "The Physiognomy of Responsibility." *Philosophy and Phenomenological Research* 82: 371-417.
- Frankfurt, H. (1971). "Free Will and the Concept of a Person." *Journal of Philosophy* 68: 5-20.
- Glover, J. (1970). *Responsibility*. Humanities Press.
- Hieronymi, P. (2004). "The Force and Fairness of Blame." *Philosophical Perspectives* 18: 115-48.
- Kane, R. (1996). *The Significance of Free Will*. Oxford University Press.
- King, M. (2012). "Moral Responsibility and Merit." *Journal of Ethics and Social Philosophy* 6(2): 1-17
- Levy, N. (2005). "The Good, the Bad and the Blameworthy." *Journal of Ethics & Social Philosophy* 1(2): 1-16.

⁵¹ I'd like to thank John Martin Fischer, Michael McKenna, Mark Schroeder, Angela Smith, and Gary Watson. Versions of this paper were given at the University of Southern California and the University of California-Riverside. My thanks go to both audiences.

- Levy, N. (2007). "Review of *Four Views on Free Will*." *Metapsychology Online Reviews* 11.40:
http://metapsychology.mentalhelp.net/poc/view_doc.php?type=book&id=3851&cn=394
- McKenna, M. (2009). "Compatibilism & Desert: Critical Comments on *Four Views on Free Will*." *Philosophical Studies* 144: 3-13.
- McKenna, M. (2012). *Conversation and Responsibility*. Oxford University Press.
- McLeod, O. (2008). "Desert." In Zalta, N (ed.), *The Stanford Encyclopedia of Philosophy (Winter 2008 Edition)*.
<http://plato.stanford.edu/archives/win2008/entries/desert/>.
- Nelkin, D. (2011). *Making Sense of Freedom and Responsibility*. Oxford University Press.
- Pereboom, D. (2007). "Hard Incompatibilism." In Fischer, J., Kane, R., Pereboom, D., and Vargas, M. *Four Views on Free Will*. Blackwell, 85-126.
- Pereboom, D. (2009). "Hard Incompatibilism and its Rivals." *Philosophical Studies* 144: 21-33
- Rosen, Gideon. (2002). Culpability and Ignorance. *Proceedings of the Aristotelian Society* 103: 61-84.
- Schroeder, M. (2010). "Value and the Right Kinds of Reason." *Oxford Studies in Metaethics* 5: 25-55.
- Shoemaker, D. (2011). "Attributability, Answerability, and Accountability: Toward a Wider Theory of Moral Responsibility." *Ethics* 121(3): 602-32.
- Smith, A. (2007). "On Being and Holding Responsible." *Journal of Ethics* 11: 465-84.
- Smith, A. (2008). "Control, Responsibility, and Moral Assessment." *Philosophical Studies* 138: 367-92.
- Smith, A. (2012). "Attributability, Answerability, and Accountability: In Defense of a Unified Account." *Ethics* 122(3): 575-89.
- Strawson, G. (1994). "The Impossibility of Moral Responsibility." *Philosophical Studies* 75:5-24.
- Strawson, P. F. (1962). "Freedom and Resentment." *Proceedings of the British Academy* 68: 1-25.
- Wallace, R. J. (1994). *Responsibility and the Moral Sentiments*. Harvard University Press.
- Watson, G. (1987). "Responsibility and the Limits of Evil: Variations on a Strawsonian Theme." In Schoeman, F. (ed.), *Responsibility, Character, and the Emotions*. Cambridge University Press.
- Watson, G. (1975). "Free Agency." *Journal of Philosophy* 72: 205-20.
- Watson, G. (1996). "Two Faces of Responsibility." *Philosophical Topics* 24: 227-48.
- Wolf, Susan. (1990). *Freedom Within Reason*. Oxford University Press.